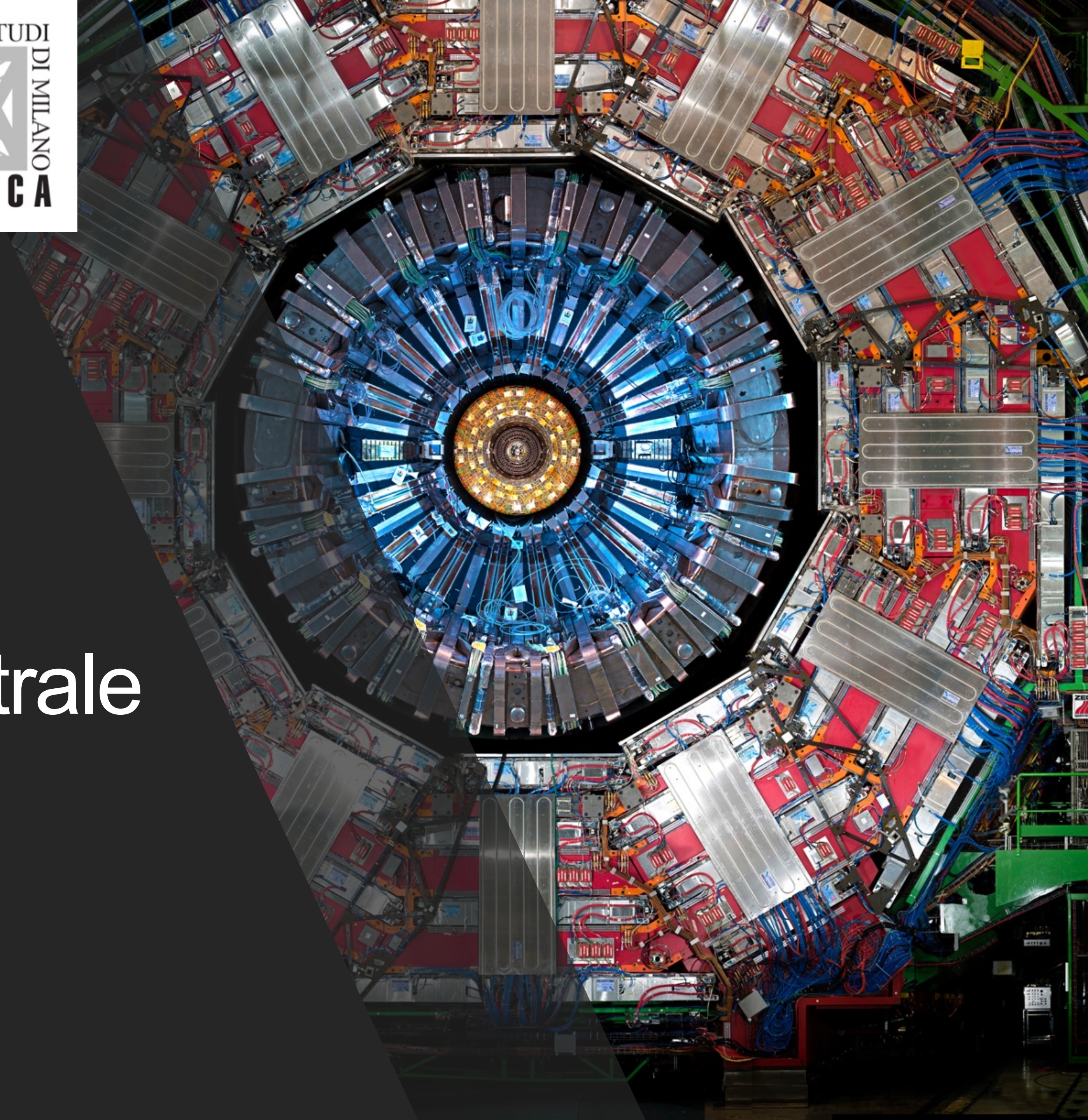




# Analisi dati a CMS

argomenti di tesi magistrale

Luca Guzzi  
14 luglio 2022

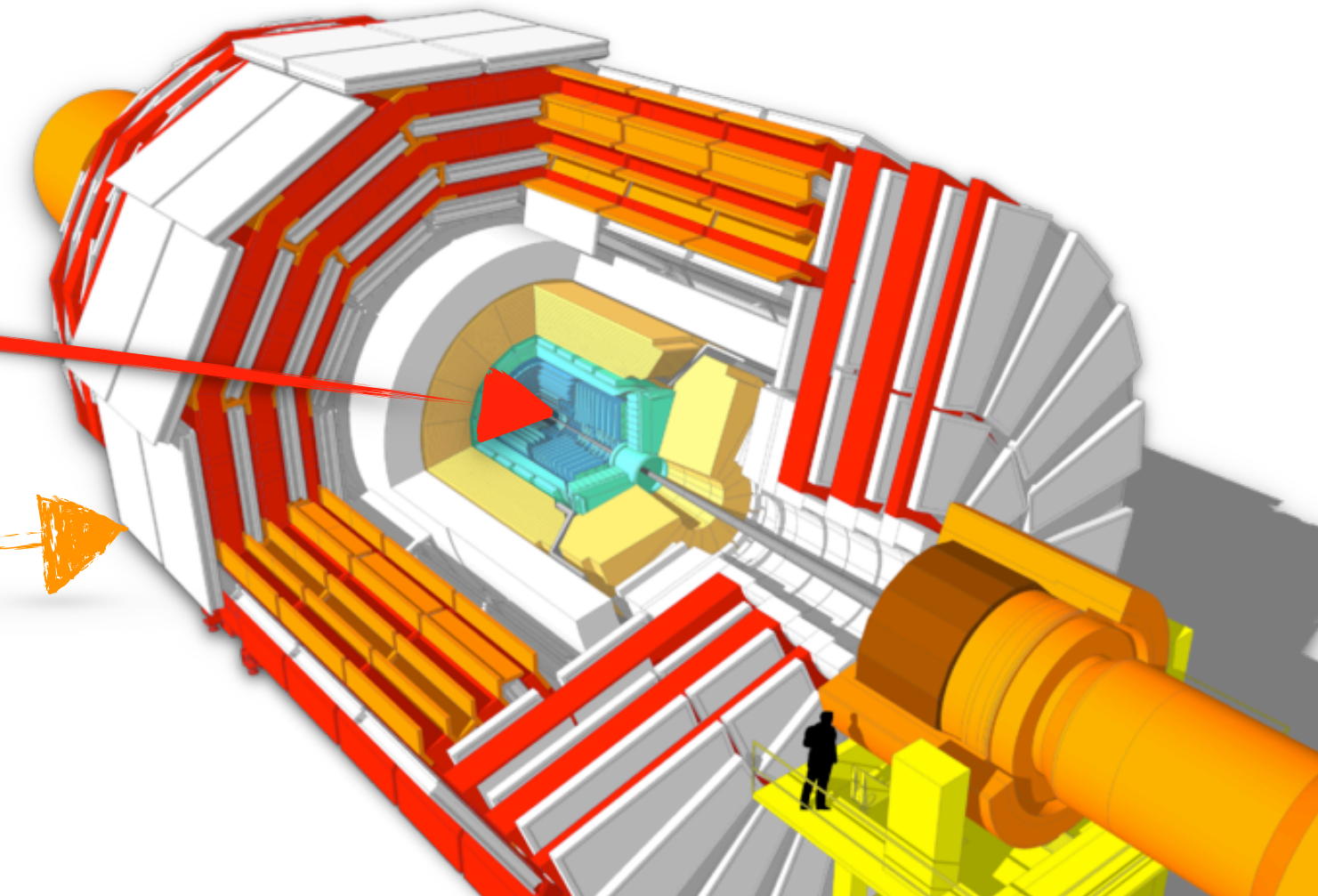


# Compact Muon Solenoid



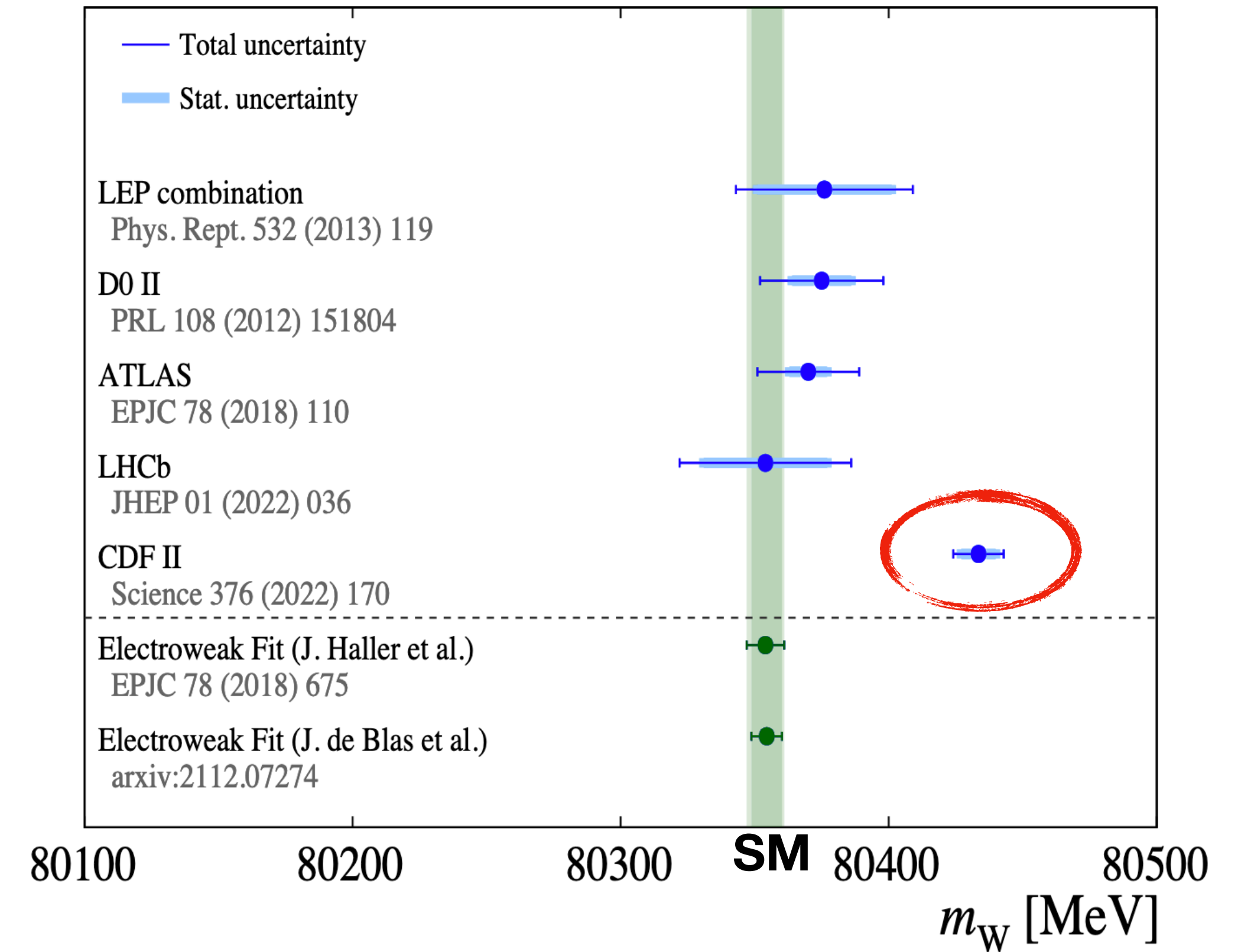
- Uno dei quattro grandi esperimenti a LHC
  - ALICE, ATLAS, CMS, LHCb
- Situato a P5 dell'anello di LHC (Cessy, FR), 100 m sotto terra
- Cilindro di 20 m di lunghezza, 7m di raggio, 12 ktons
- magnete superconduttore da 3.8 T
- In grado di *misurare e ricostruire ermeticamente* le particelle prodotte dalle collisioni di **protoni**

**pp collision!** una ogni 25 ns, 40 M al secondo!



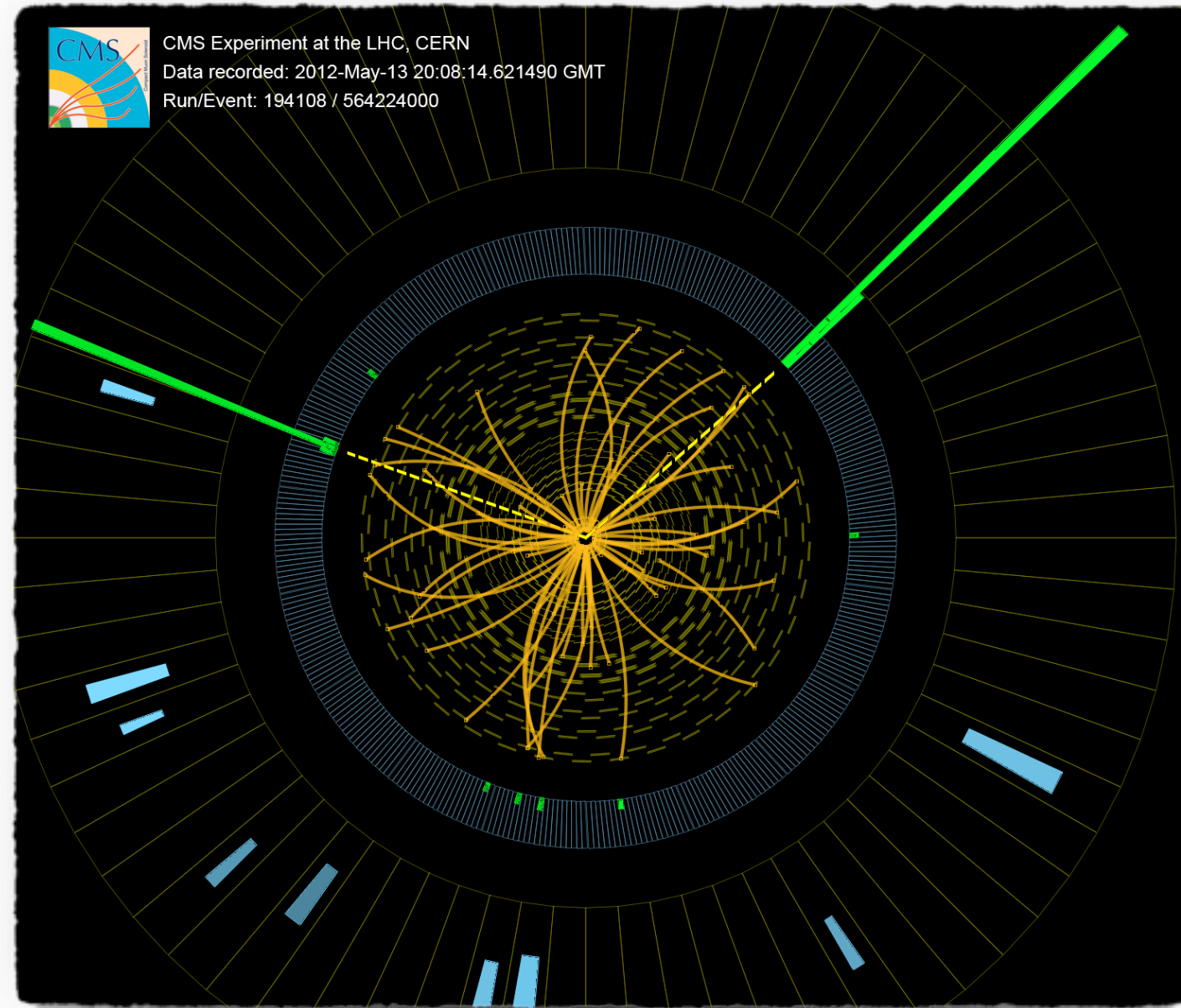
# Modello Standard e Oltre

- Nel 2012 ATLAS e CMS hanno confermato l'esistenza di una particella compatibile con il bosone di Higgs. E' tutto qui?
  - Conosciamo tutte le sue proprietà? (No)
- Cosa genera la gerarchia di massa delle particelle? Perché l'Higgs è così leggero? Perché il top è così pesante? Esiste una simmetria tra le famiglie di quark e di leptoni? Cos'è la materia oscura? Il MS è valido anche a energie molto elevate? ...?



Il recente risultato di CDF-II (Tevatron @Fermilab) sulla stima della mass del bosone **W** presenta una chiara tensione con il MS

# Analisi dati in breve

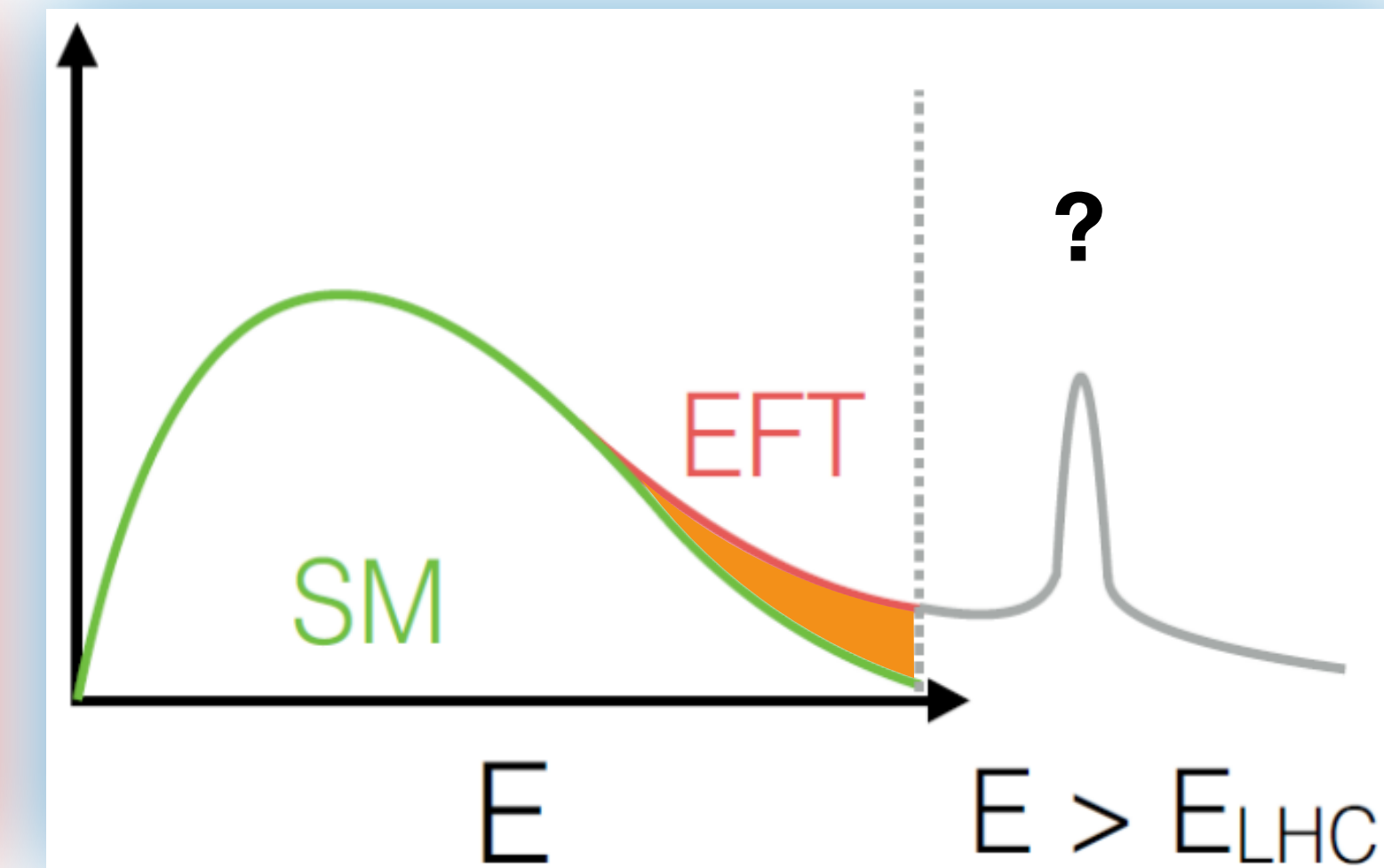
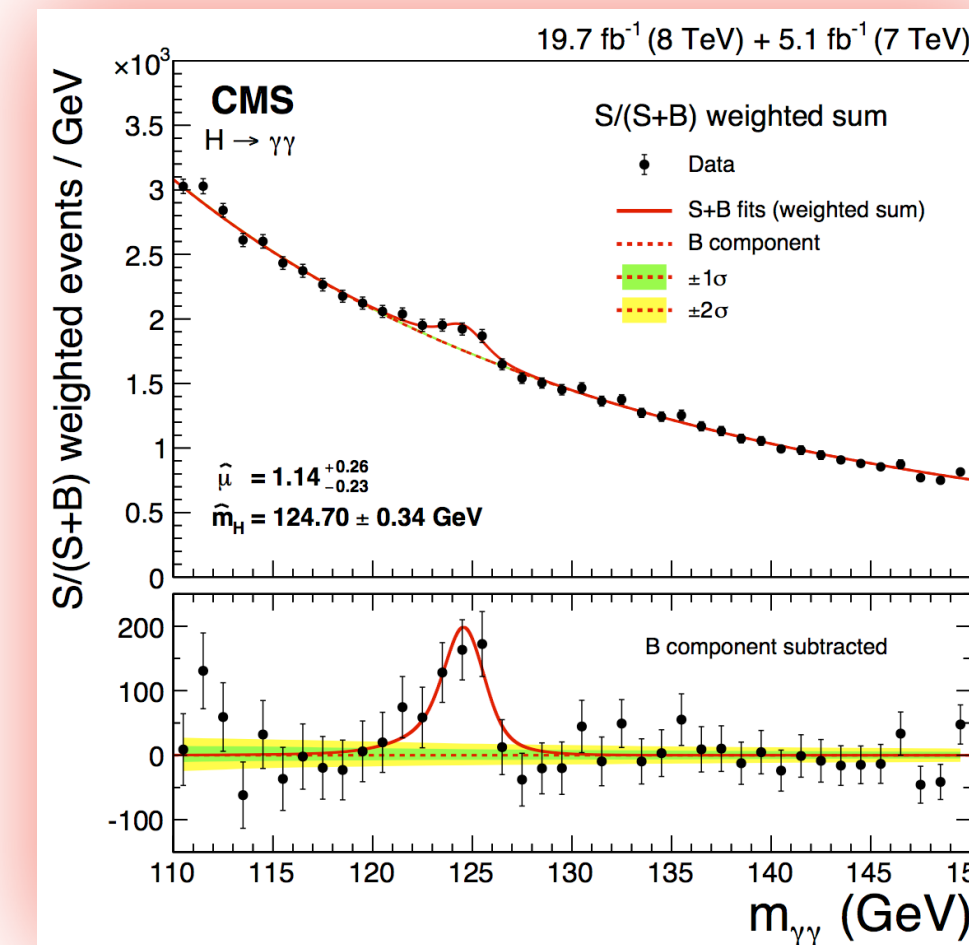


**evento candidato  
Higgs  $\rightarrow \gamma\gamma$**   
sono chiaramente visibili i depositi di energia di due fotoni nel calorimetro elettromagnetico (barre verdi)

- I vari sotto-detector generano un segnale al transito delle particelle
- Segnale ricostruito tramite impiego di svariati algoritmi
- Stato finale di un decadimento individuato selezionando gli eventi ricostruiti
  - confronto con modelli vari (Modello Standard, modelli di nuova fisica)

Possibili scenari di ricerca:

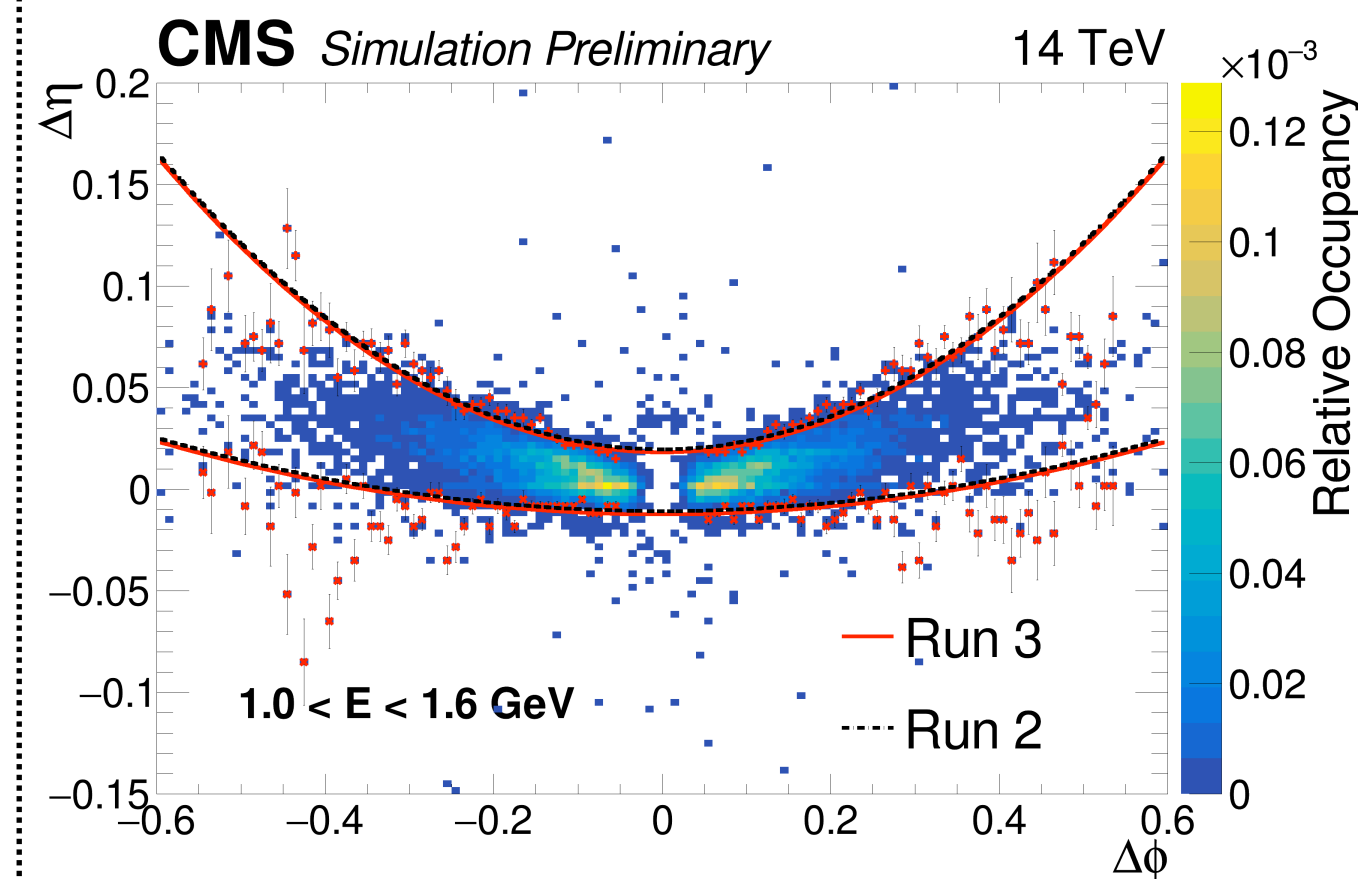
- evidenza **diretta** di nuova fisica (nuove particelle)
- evidenza **indiretta** (risultati diversi dalle previsioni) come deviazione dal MS
  - possibile interpretazione dei risultati come *Teorie di campo efficaci (EFT)*: estensione del MS tramite operatori "efficaci" che approssimano la dinamica di modelli più generali validi a scale di energia più grandi (es. *storici*: il pione media l'interazione nucleare forte, teoria elettrodebole di Fermi)



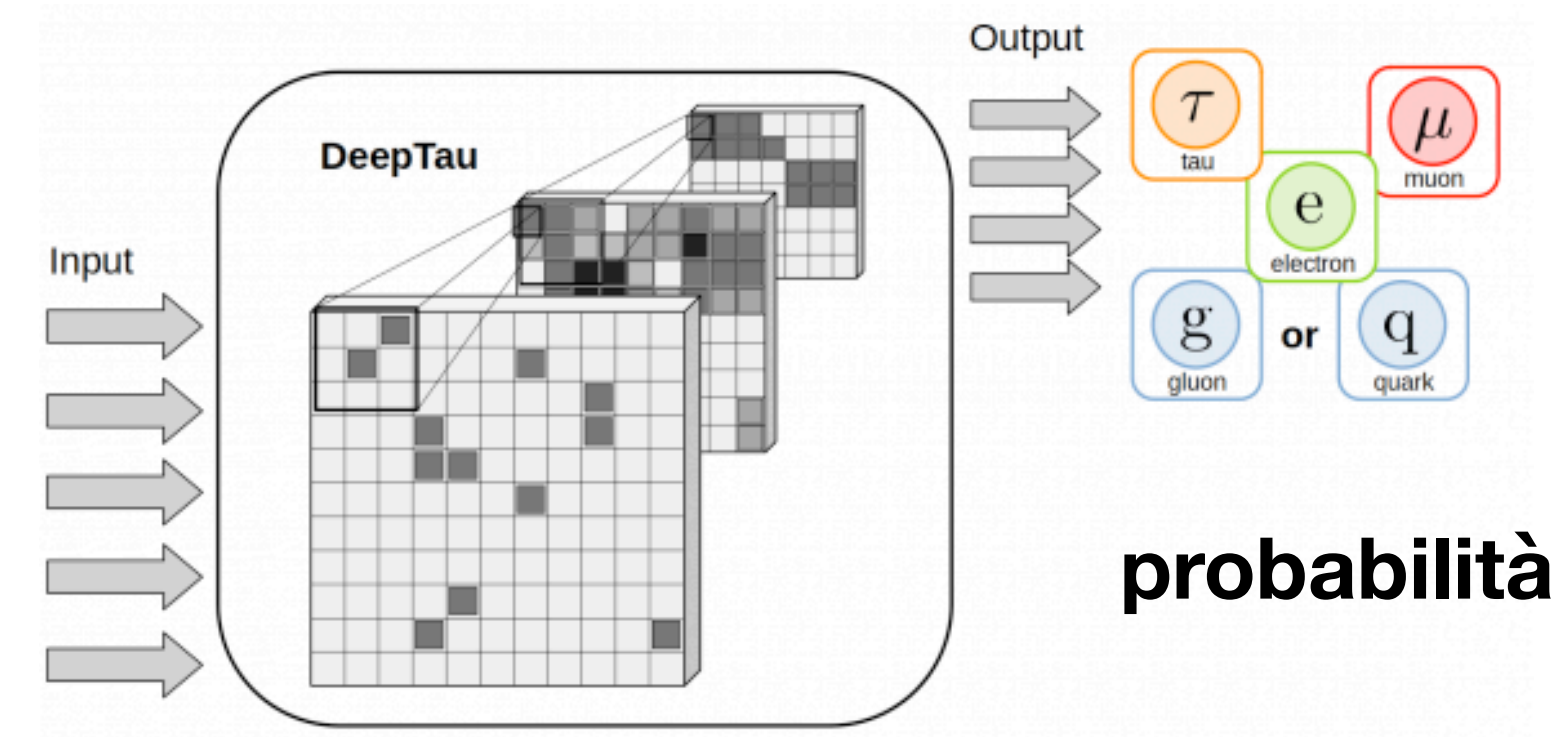
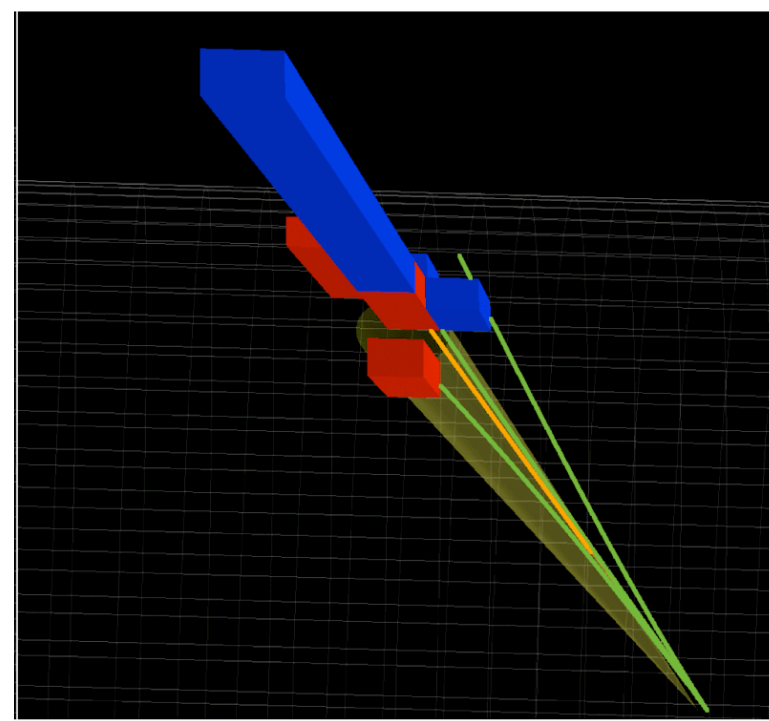
# Le opportunità in HEP

- utilizzo di **tecniche computazionali e statistiche avanzate** per l'analisi di ingenti moli di dati
- utilizzo di **tecniche di regressione, inferenza e ricostruzione moderne** (neural network e machine learning) nei campi più disparati (classificazione di eventi, ricostruzione di segnali, sviluppo di algoritmi di selezione -trigger-, confronto di modelli e ricerca di segnali di nuova fisica)

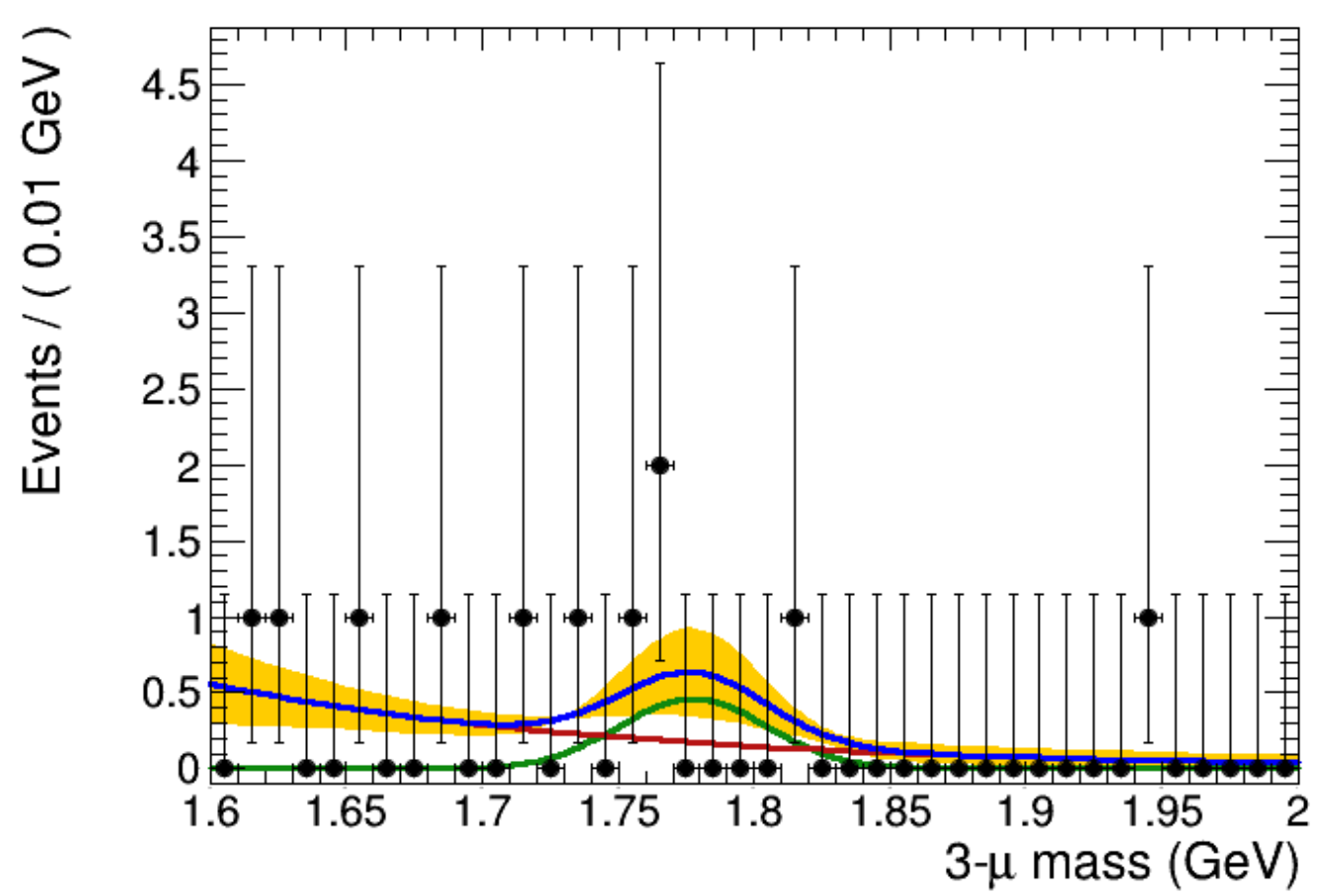
L'algoritmo di clustering del calorimetro elettromagnetico di CMS seleziona cluster di cristalli in maniera puramente geometrica. Si può fare meglio con ML?



## rete convolutiva



qual è la probabilità che questo "getto" di particelle sia un evento tau→pioni?



Come possiamo quantificare l'eccesso di eventi osservato? E' significativo?

# Le opportunità in HEP

---

- **ambiente di lavoro internazionale stimolante**, possibilità di **stage all'estero** (e.g. CERN, FermiLab, DESY, LLR, ...)
- possibilità di **contribuire in modo rilevante** ad analisi complesse e di grande impatto
- capacità acquisibili (per HEP e non solo!)
  - capacità di risoluzione di problemi complessi
  - utilizzo di **linguaggi di programmazione moderni** (e.g. C++ and python)
  - utilizzo di **software di manipolazione dati** (ROOT, pandas, numpy, scipy, ...), **analisi dati** e **machine learning** (ROOT, tensorflow, pytorch, keras)

# Argomenti di tesi

---

1. Application of state-of-the-art artificial intelligence algorithms to the **Higgs boson self-coupling measurement**
2. A sensitivity **study of the production of two Higgs processes to dimension-6 EFT** operators at the LHC.
3. Search for **Flavour Changing Neutral Current decays of B0 mesons** with machine-learning-based analysis techniques
4. Implementation of Machine Learning techniques in the **Vector Boson Scattering /Fusion WW and WV production channel** to search beyond the standard model physics in the effective field theory (EFT) approach.
5. Search for beyond the standard model physics through the phenomenological **study of effective field theory (EFT) CP-violating operators** of dimension 6 to understand the matter anti-matter asymmetry of the universe.
6. **Anomaly Detection** methods for high energy physics processes
7. **Machine Learning for non-prompt leptons** for VBS studies

<https://www.fisica.unimib.it/it/didattica/corsi-studio/corso-laurea-magistrale-fisica/argomenti-prova-finale-della-laurea-magistrale/argomenti-prova-finale-della-laurea-magistrale-fisica-delle-particelle>

# Produzione di coppie di Higgs:

## $HH \rightarrow bb \tau\tau$

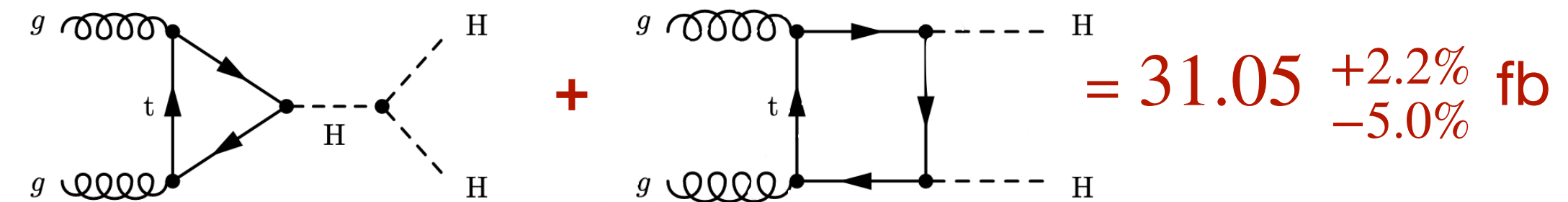
- 2012: ATLAS e CMS annunciano la scoperta di un bosone neutro di massa 125 GeV
  - proprietà del bosone compatibili con le previsioni del MS

- misura dell'accoppiamento trilineare **HHH** su dati **CMS**

- come interagisce con se stesso?
- possiamo conoscere il suo potenziale con precisione maggiore?
- Il rate di produzione in collisioni pp è quello atteso?
  - ricerca di **fisica oltre il MS** e nuove particelle (Radioni, Gravitoni, ...)

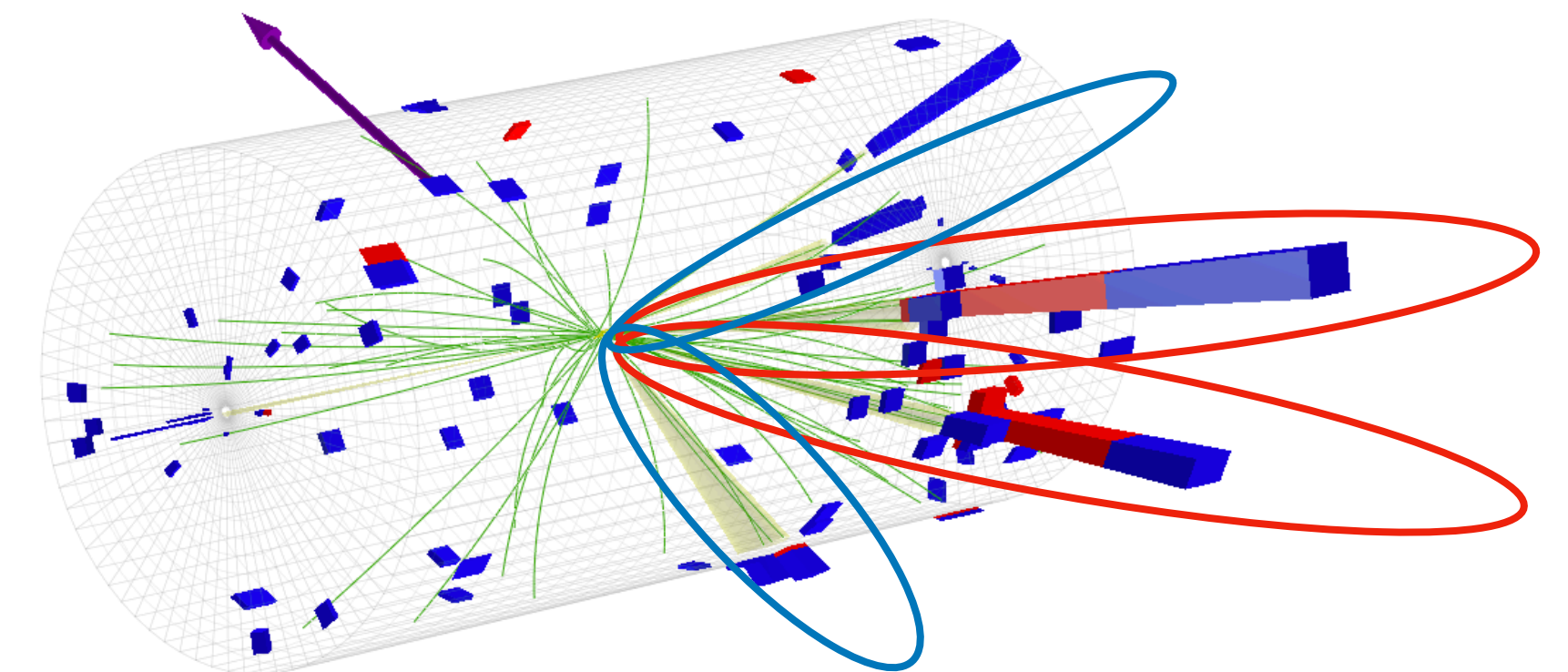
- sezione d'urto molto piccola (1/1000 rispetto a singolo Higgs)
  - utilizzo di tecniche di Machine Learning per ridurre il fondo e selezionare efficacemente gli eventi
  - utilizzo di tecniche di Machine Learning per la stima di quantità utili alla selezione del segnale (massa  $\tau\tau$ , identificazione b-jet)

**theory:**



CMS Experiment at LHC, CERN  
 Data recorded: Wed Oct 3 11:09:52 2018 UTC  
 Run/Event: 323954 / 16341342  
 Lumi section: 9  
 Orbit/Crossing: 2209447 / 3295

**experiment:**



$HH \rightarrow bb \tau\tau?$

Come e quanto possiamo esserne sicuri?  
 Ogni quanto vediamo questo tipo di evento?

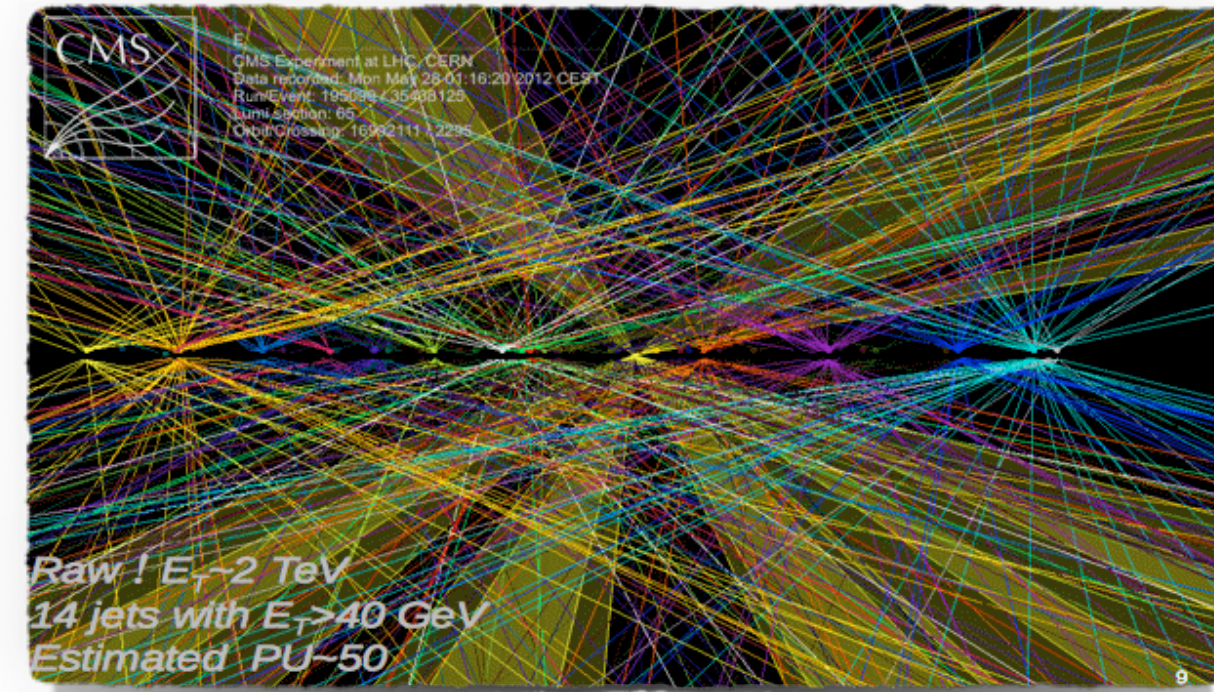


[tommaso.tabarelli@unimib.it](mailto:tommaso.tabarelli@unimib.it)  
[alessio.ghezzi@mib.infn.it](mailto:alessio.ghezzi@mib.infn.it)  
[federico.deguio@unimib.it](mailto:federico.deguio@unimib.it)  
[marco.lucchini@unimib.it](mailto:marco.lucchini@unimib.it)  
[martina.malberti1@unimib.it](mailto:martina.malberti1@unimib.it)  
[andreadavide.benaglia@unimib.it](mailto:andreadavide.benaglia@unimib.it)  
[pietro.govoni@mib.infn.it](mailto:pietro.govoni@mib.infn.it)  
[marco.paganoni@mib.infn.it](mailto:marco.paganoni@mib.infn.it)  
 (parton-level studies)

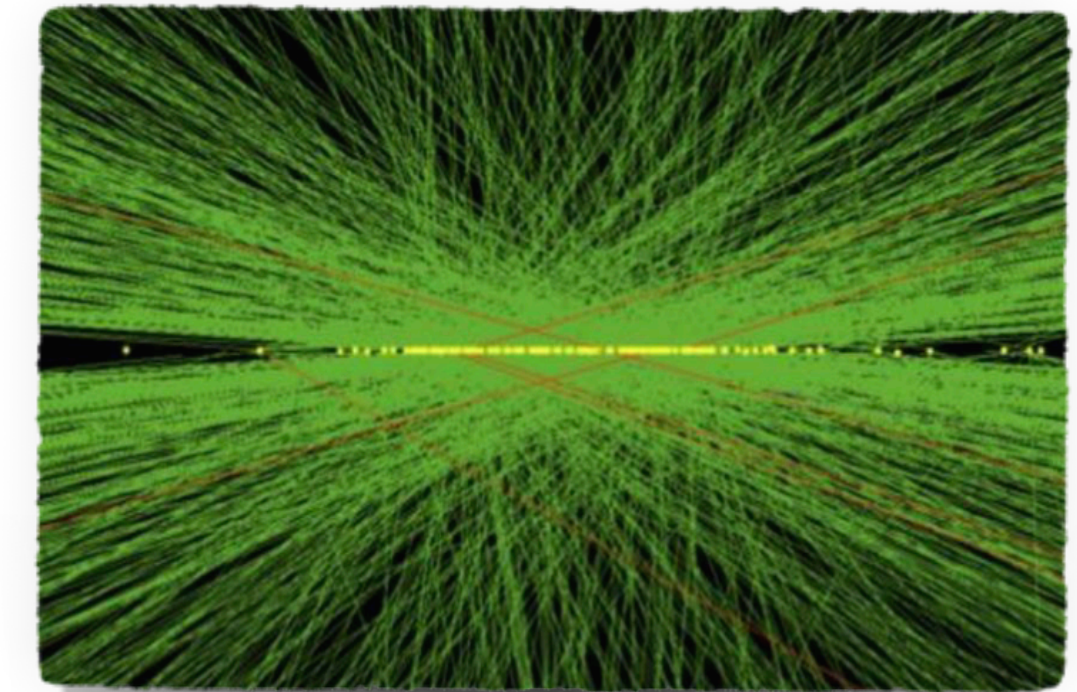
# Produzione di coppie di Higgs:

## $HH \rightarrow bb \gamma\gamma$

- LHC fase 2 (dal 2029) porterà il pileup a circa 200 interazioni pp medie per collisione (attuali 30)
  - fondamentale ridurre il pileup
  - MTD in sviluppo per fase 2 anche a Milano Bicocca
  - più dettagli nella prossima presentazione
- MTD: MIP timing layer
  - ricostruzione dei vertici di interazione ora basata su informazione spaziale (posizione delle tracce propagate verso il punto di interazione)
- **Possibile miglioramento di alcune analisi, come produzione di coppie di Higgs**
- Studio della sensibilità delle analisi HH agli operatori efficaci (EFT) a **livello partonico**

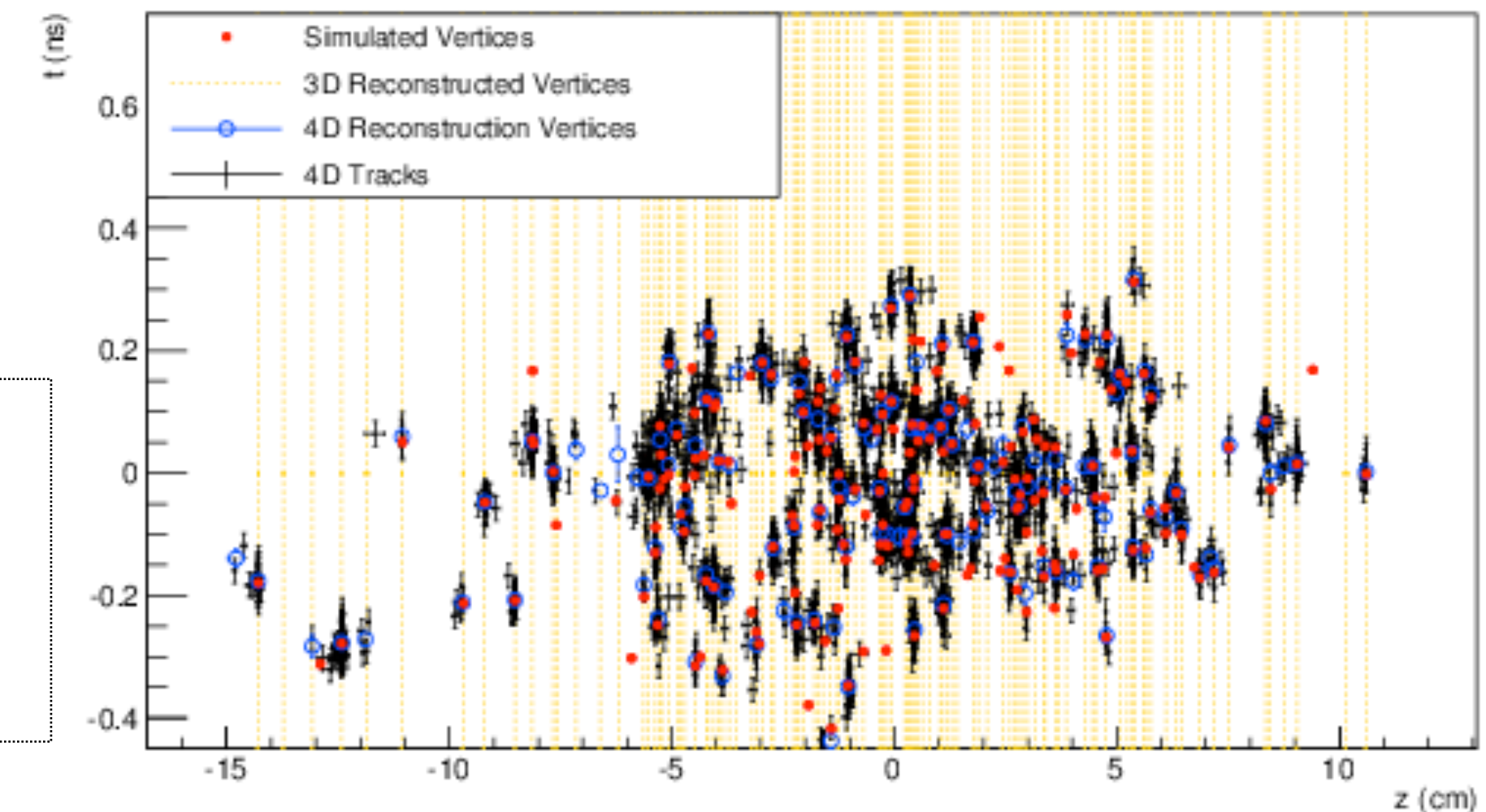


evento del 2012 (50 vertici)



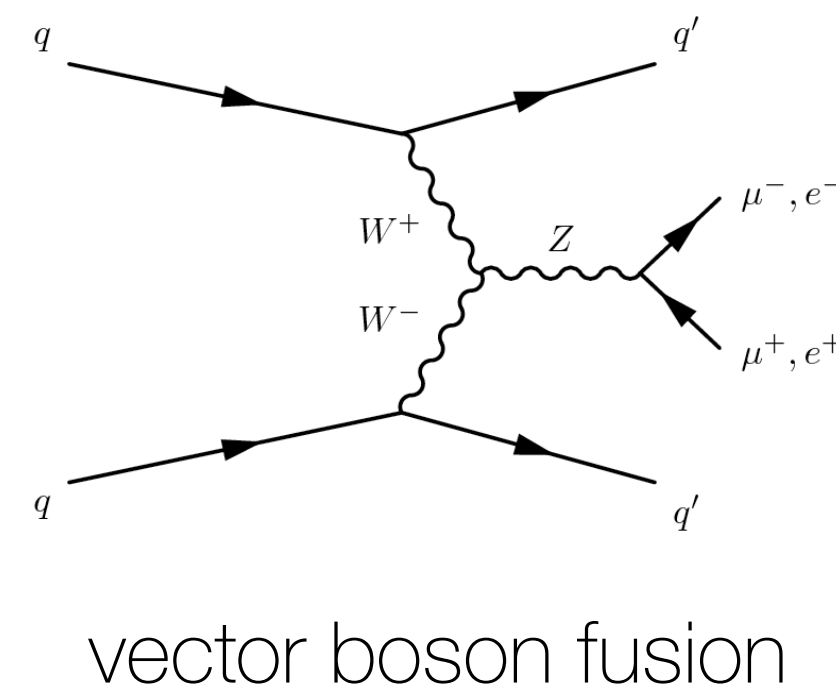
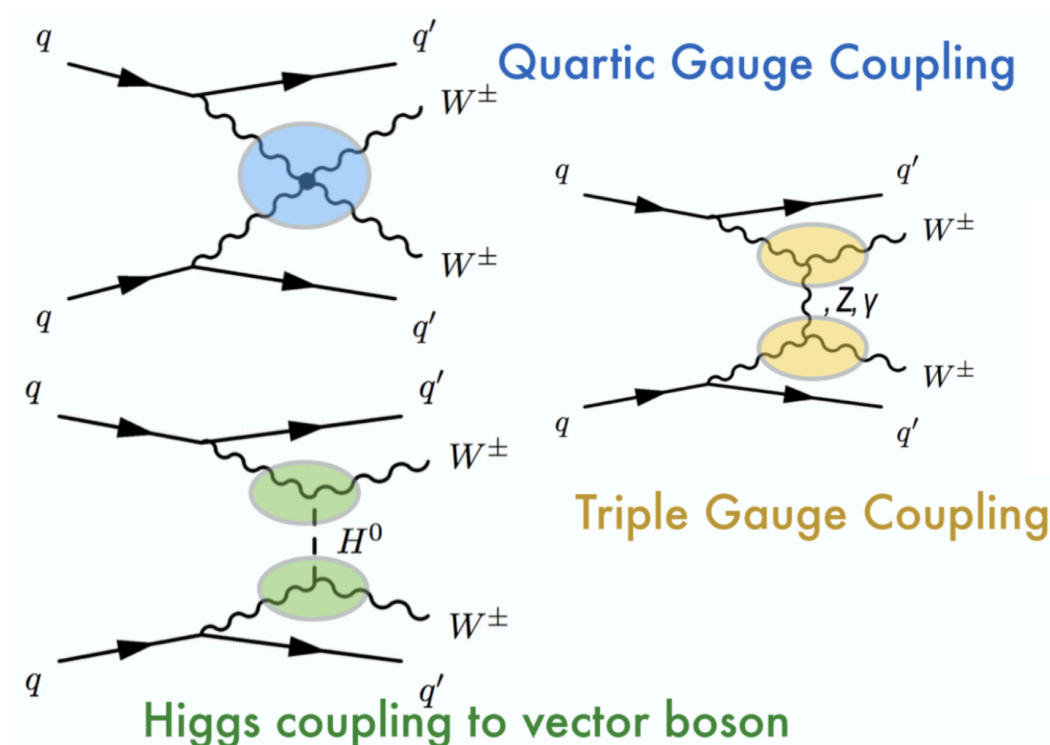
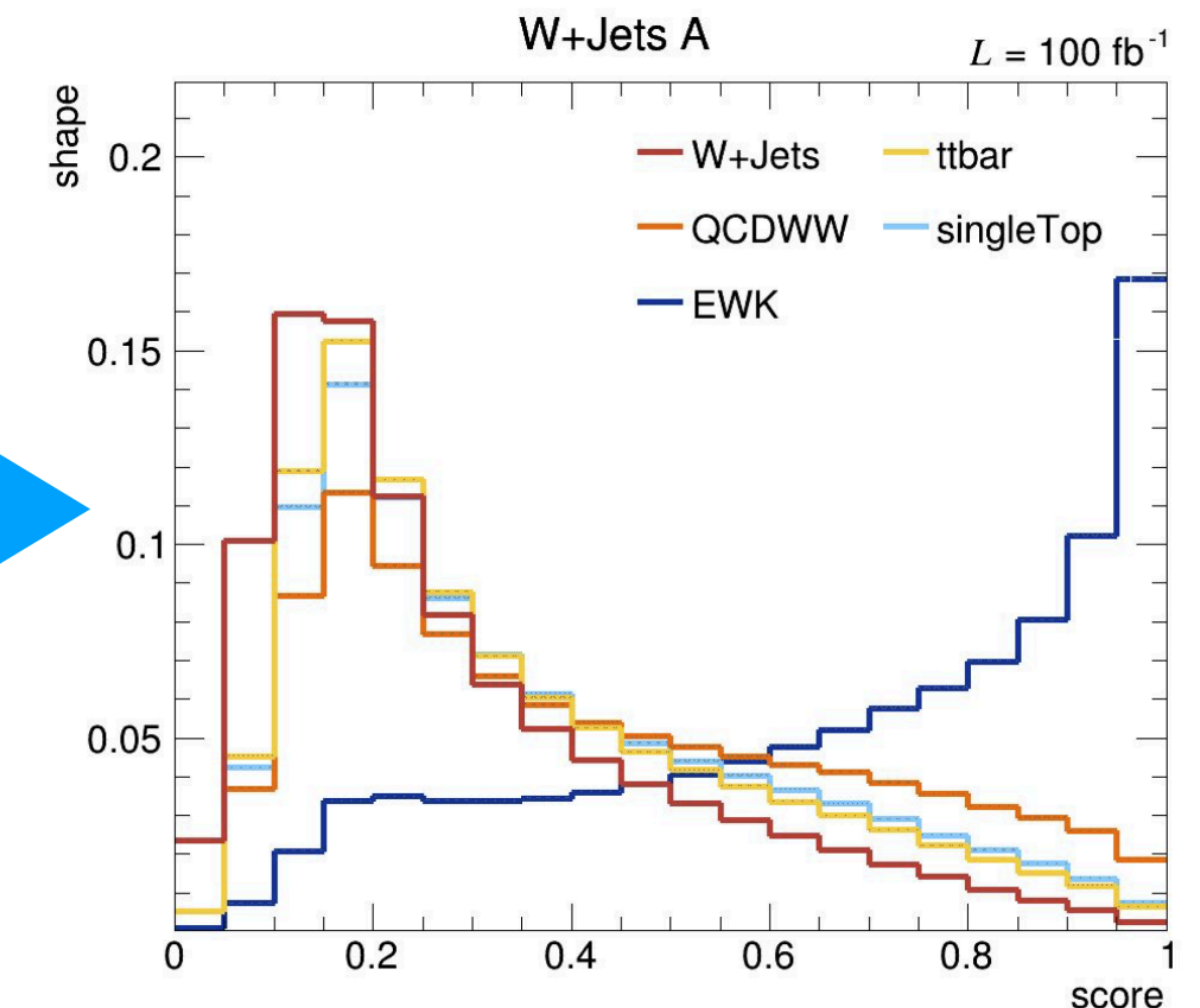
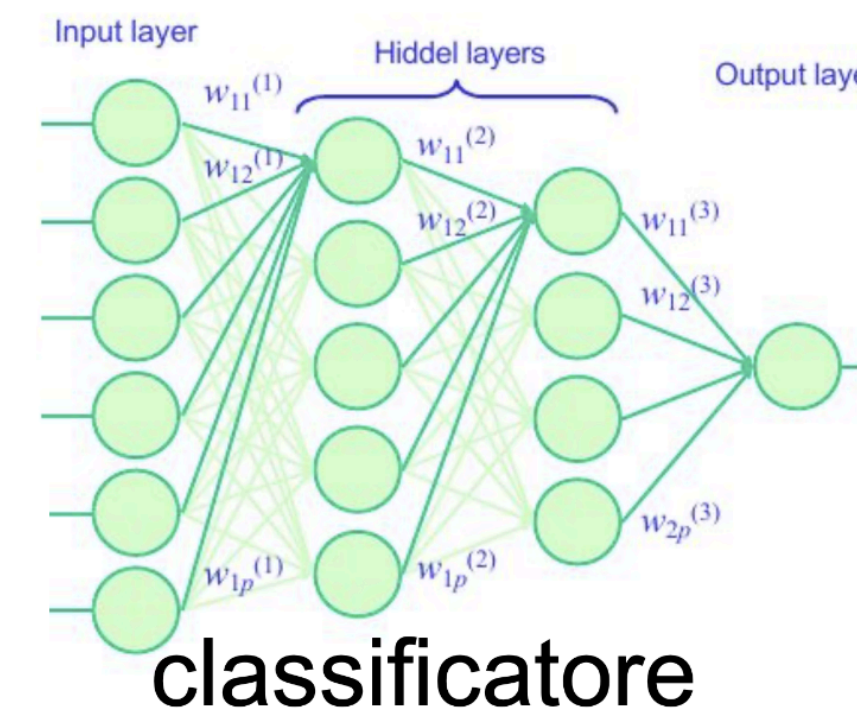
evento a PU 140

**MTD** aggiunge l'informazione temporale a quella spaziale



# Vector Boson Scattering / Fusion

- Produzione di coppie di bosoni vettori in interazioni pp
  - **test del settore elettrodebole del MS**
    - molti contributi alla produzione: molte possibilità di deviazioni dal MS
  - analisi MS consolidata: osservazione di produzione WW a  $>5$  sigma!
  - cosa fare ora? **test di validità del MS e studio di EFT** con analisi di precisione
  - differenti stati finali in esame:  $W^+W^+$ ,  $W^-W^+$ ,  $WZ$ 
    - **combinazione dei canali** interessante sia per verifica MS che per studi EFT

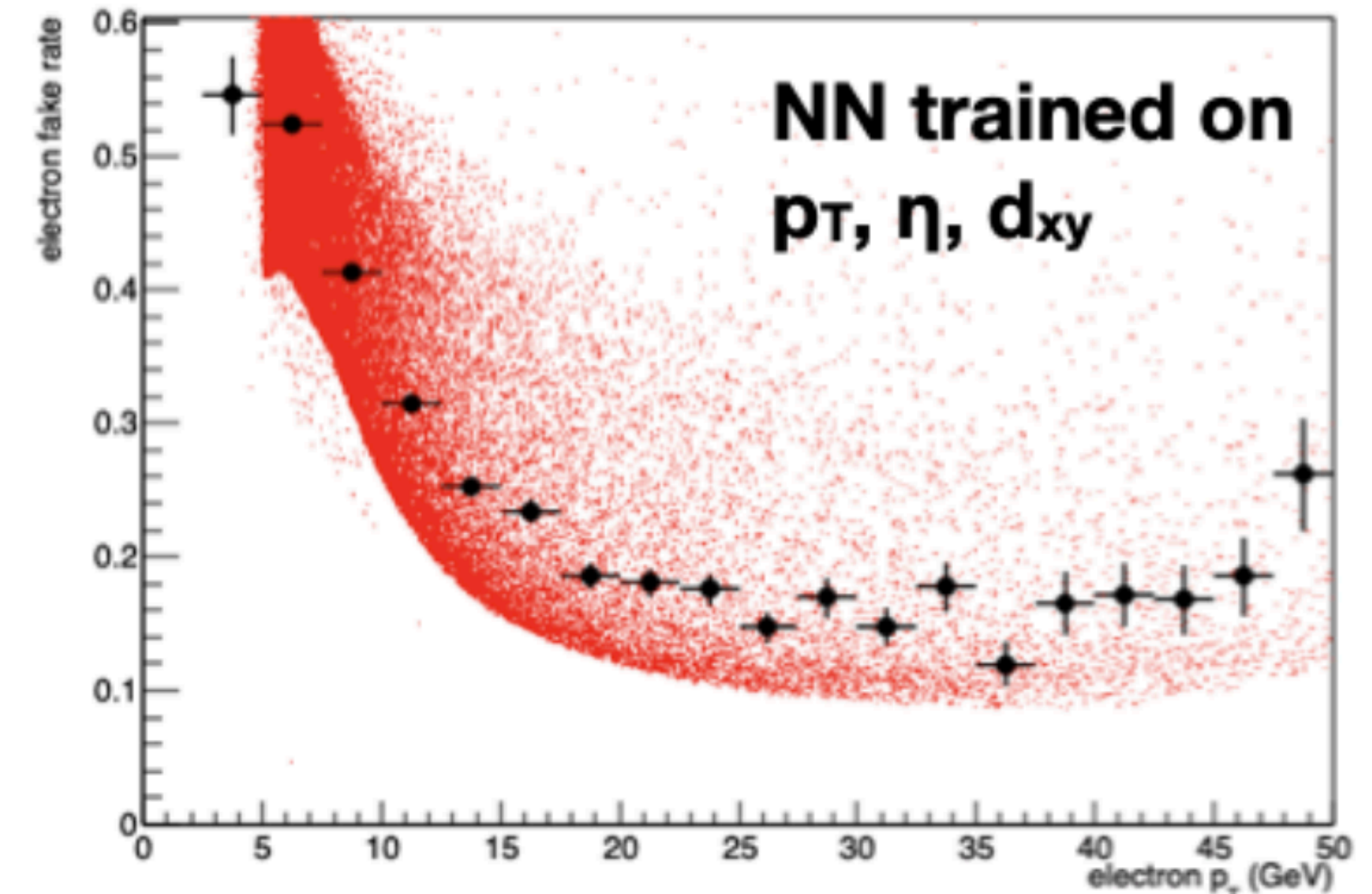


- VBS e VBF sono **processi rari**: necessario ridurre il fondo
  - utilizzo di Deep Neural Network per la classificazione degli eventi
  - utilizzo di Deep Neural Network per la soppressione del fondo di leptoni (non-prompt lepton suppression)

vector boson scattering

# VBS/F: non-prompt leptons

- Presenza di leptoni carichi (muoni, elettroni) nello stato finale VBS (e.g.  $W \rightarrow \mu\nu$ )
  - i jet identificati erroneamente come leptoni peggiorano il risultato dell'analisi (**fake-leptons** o non-prompt leptons)
- Possibili correzioni ricavate dai dati (esempio, studiando eventi  $Z \rightarrow \mu\mu$ )
  - fake-probability come funzione discreta di due variabili: momento del fake-lepton e pseudorapidità
- **Machine Learning** permetterebbe di ricostruire la dipendenza non binnata del fake rate in funzione di queste
  - introducendo più informazioni (più variabili) si otterrebbero risultati più accurati

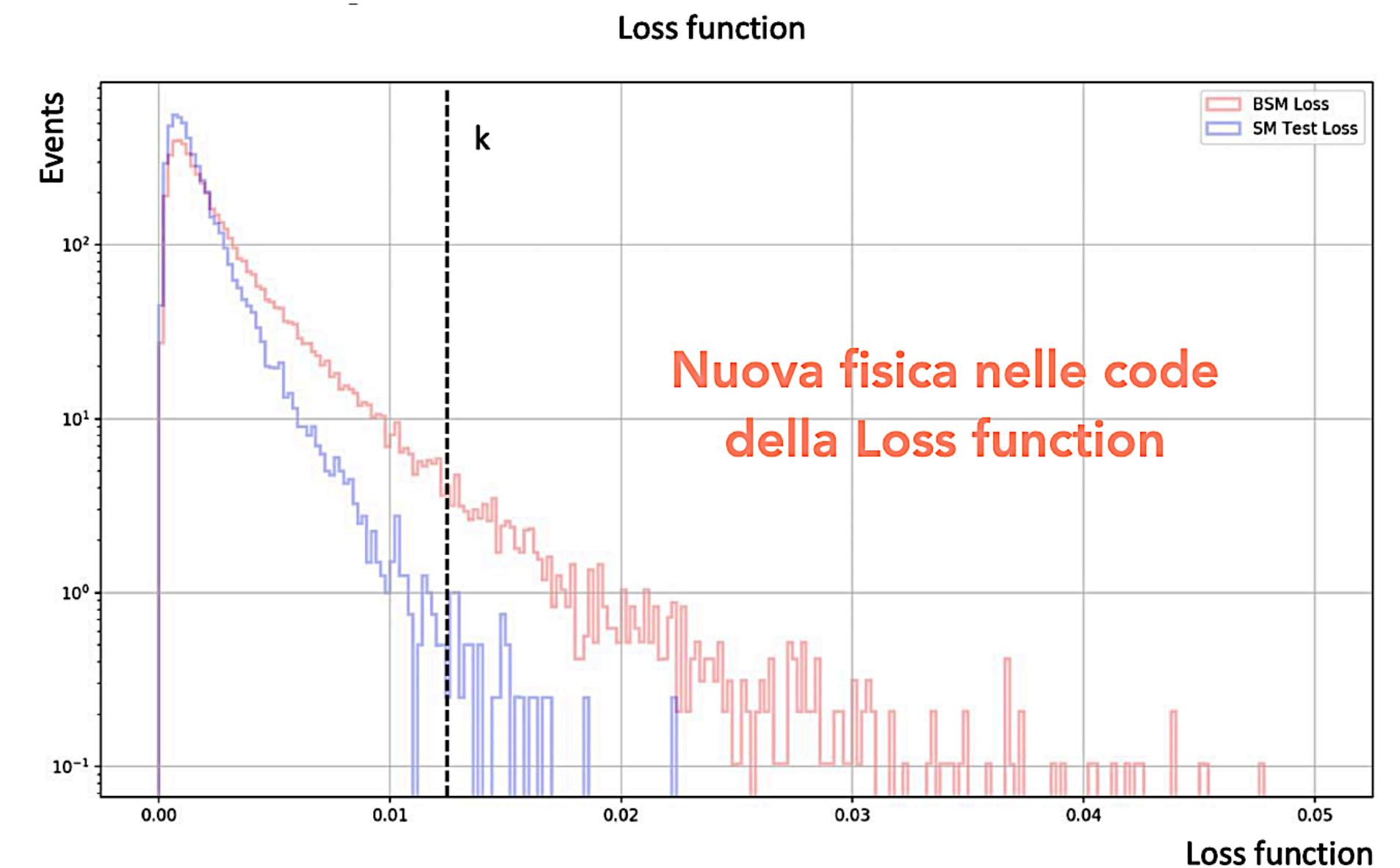
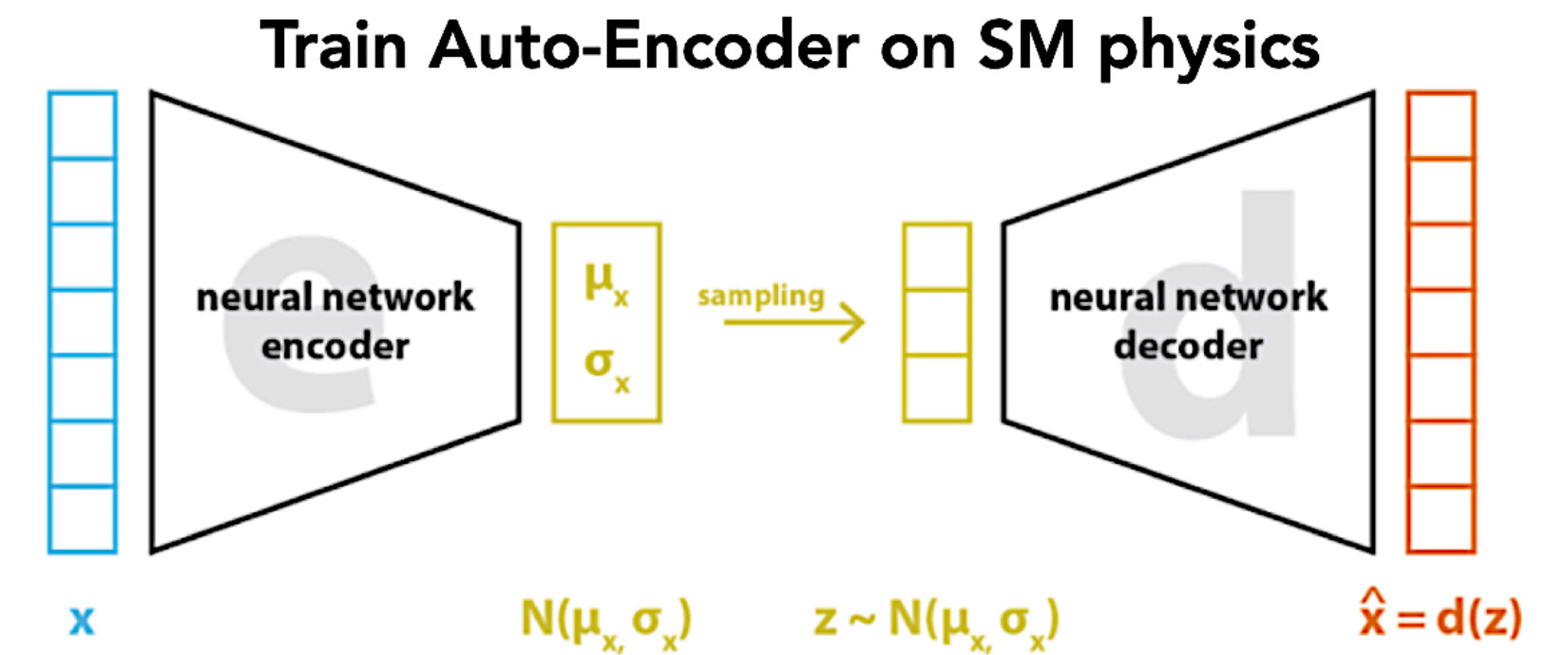


mis-ID ottenuto da tecniche ML (punti rossi) o con tecniche più classiche (punti neri) in funzione del  $p_T$  degli elettroni (esempio per un'altra analisi CMS - ricerca di neutrini pesanti)

- **ML** riproduce le tecniche classiche in media
- **ML** introduce informazione aggiuntiva (dispersione attorno alla media)
- possiamo farlo anche per VBS!

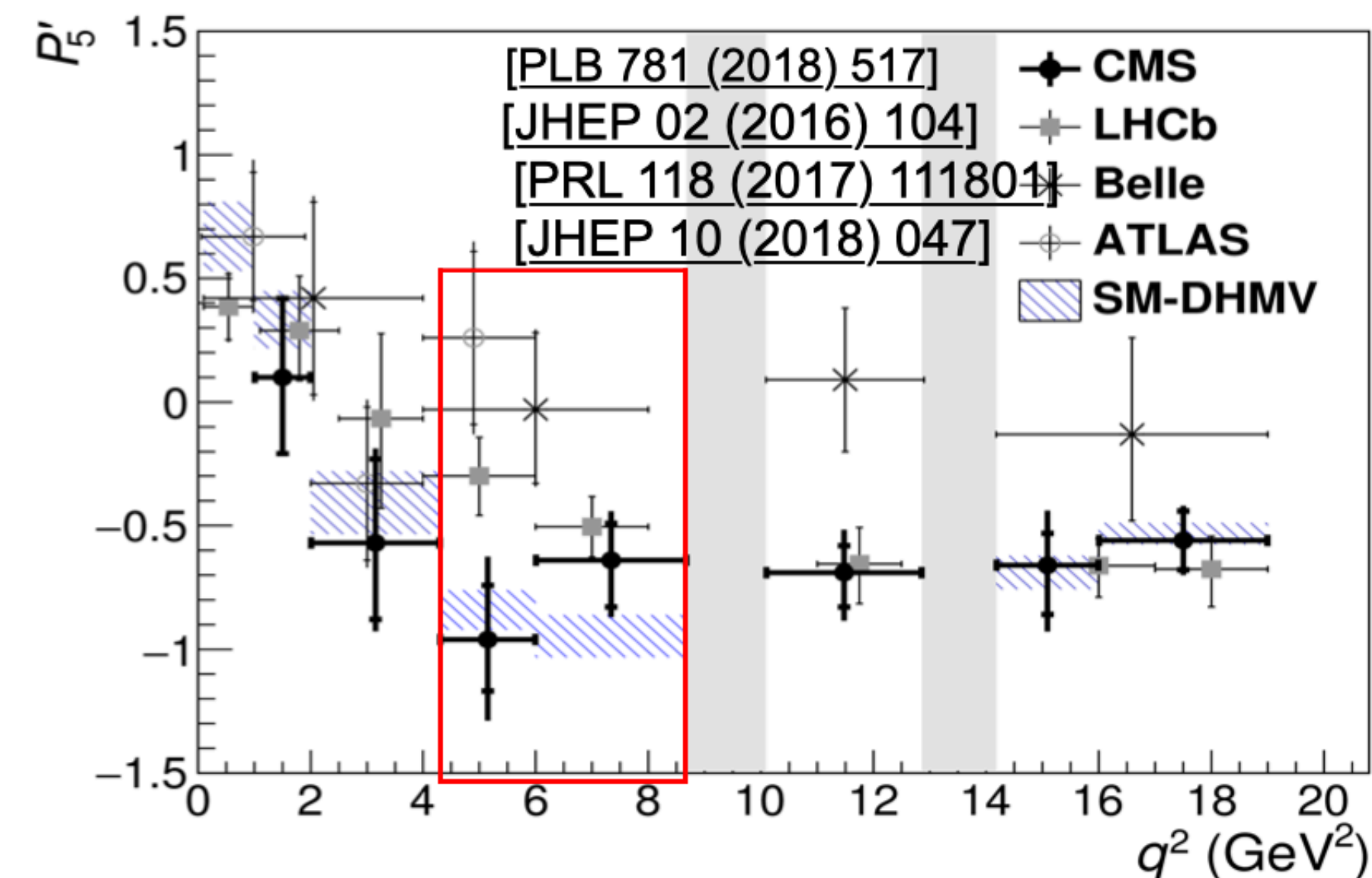
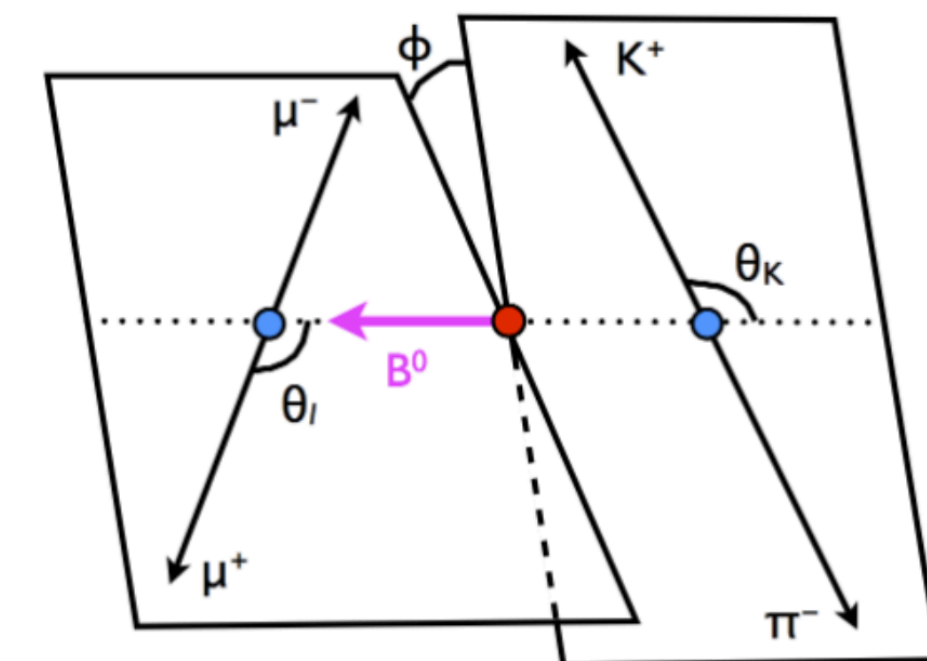
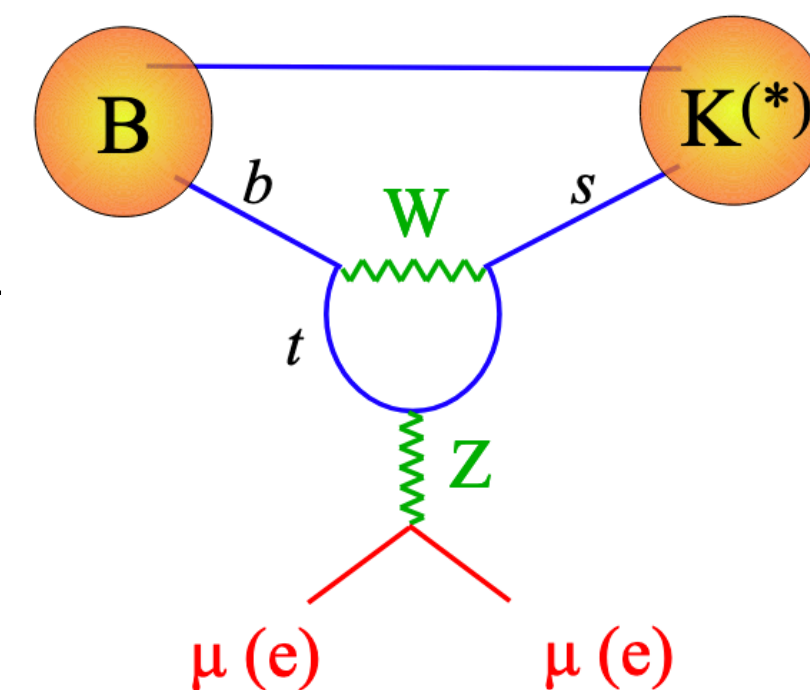
# VBS/F: anomaly detection

- **Auto-encoders:** tecnica di *unsupervised machine learning* adattata al riconoscimento di "ciò che non sembra MS" (anomaly detection)
  - approccio alternativo ad un'analisi mirata
  - possibilità di isolare eventi che si differenziano in modo significativo dal MS
  - ricerca di deviazioni dal MS di quanto viene osservato e test di compatibilità con le previsioni EFT



# $b \rightarrow s$ nei decadimenti del $B^0$

- Tensioni con il MS nei decadimenti  $b \rightarrow s$  osservate da LHCb
  - importante il feedback di molti esperimenti
  - CMS può contribuire alle analisi a basso momento trasverso e bassa massa (fisica dei mesoni B e D)
- Fisica di precisione per lo studio del decadimento  $B^0 \rightarrow K^*(K \pi) \mu\mu$ 
  - analisi angolare 4-dim del decadimento e confronto dei parametri che lo descrivono con le previsioni del MS
    - ▶ Ricerca di possibili deviazioni dal MS e quindi nuova fisica
- Tecniche di machine learning per la selezione del segnale
- Utilizzo di GPU per la parametrizzazione 3-dim dell'efficienza di selezione di segnale e fondo



# Contatti

## HH $\rightarrow$ bb $\tau\tau$ (doppio-Higgs su dati CMS)

- Mauro Dinardo [mauro.dinardo@mib.infn.it](mailto:mauro.dinardo@mib.infn.it)
- Sandra Malvezzi [sandra.malvezzi@mib.infn.it](mailto:sandra.malvezzi@mib.infn.it)
- Simone Gennai [simone.gennai@mib.infn.it](mailto:simone.gennai@mib.infn.it)
- Paolo Dini [paolo.dini@mib.infn.it](mailto:paolo.dini@mib.infn.it)
- Luca Guzzi [luca.guzzi@mib.infn.it](mailto:luca.guzzi@mib.infn.it)

## Unsupervised ML (anomaly detection)

- Simone Gennai [simone.gennai@mib.infn.it](mailto:simone.gennai@mib.infn.it)
- Pietro Govoni [pietro.govoni@mib.infn.it](mailto:pietro.govoni@mib.infn.it)

## doppio-Higgs a livello partonico

- Pietro Govoni [pietro.govoni@mib.infn.it](mailto:pietro.govoni@mib.infn.it)
- Marco Paganoni [marco.paganoni@mib.infn.it](mailto:marco.paganoni@mib.infn.it)

## b $\rightarrow$ s ( $B^0 \rightarrow K^* \mu\mu$ )

- Paolo Dini [paolo.dini@mib.infn.it](mailto:paolo.dini@mib.infn.it)
- Mauro Dinardo [mauro.dinardo@mib.infn.it](mailto:mauro.dinardo@mib.infn.it)
- Sandra Malvezzi [sandra.malvezzi@mib.infn.it](mailto:sandra.malvezzi@mib.infn.it)

## MTD per doppio-Higgs

- Tommaso Tabarelli [tommaso.tabarelli@unimib.it](mailto:tommaso.tabarelli@unimib.it)
- Alessio Ghezzi [alessio.ghezzi@mib.infn.it](mailto:alessio.ghezzi@mib.infn.it)
- Federico De Guio [federico.deguio@unimib.it](mailto:federico.deguio@unimib.it)
- Marco Lucchini [marco.lucchini@unimib.it](mailto:marco.lucchini@unimib.it)
- Martina Malberti [martina.malberti1@unimib.it](mailto:martina.malberti1@unimib.it)
- Andrea Benaglia [andreadavide.benaglia@unimib.it](mailto:andreadavide.benaglia@unimib.it)

## Vector Boson Scattering / Fusion

- Pietro Govoni [pietro.govoni@mib.infn.it](mailto:pietro.govoni@mib.infn.it)
- Andrea Massironi [andrea.massironi@mib.infn.it](mailto:andrea.massironi@mib.infn.it)
- Marco Paganoni [marco.paganoni@mib.infn.it](mailto:marco.paganoni@mib.infn.it)
- Flavia Cetorelli [flavia.cetorelli@cern.ch](mailto:flavia.cetorelli@cern.ch)



CMS Experiment at the LHC, CERN

Data recorded: 2022-Jul-05 14:48:56.743936 GMT

Run / Event / LS: 355100 / 51596902 / 53

*Grazie per  
l'attenzione!*

